Justification for Continued Use

The rule-based framework for the MYCIN consultation program has demonstrated its applicability to medical domains. In addition, other SUMEX-AIM projects depend on our continued maintenance and development of the framework (see section on Collaboration).

Because of Dr. Shortliffe's commitment, future developments of the program will be based on clinical problems which he encounters. We have not identified the precise areas at this time, however.

Computing Requirements

We perceive no additional requirements either from the system or from outside.

Recommendations

Community Development - Create a "visiting scientist" position for individuals who want to spend a few months at Stanford learning about artificial intelligence in medicine. If it carries a stipend, or partial salary support for sabbatical leave, finding qualified individuals should not be difficult.

Resource Development - Undertake, as a long range goal, solution of the problem of making large research programs available (in some form) to practicing physicians and biomedical scientists. This would require persons who understand the complexities of these programs and of the people using them, and could involve purchase and support of small computers as well.

4.2.7 PROTEIN STRUCTURE PROJECT

Protein Structure Modeling Project

Prof. E. Feigenbaum and Dr. R. Engelmore (Computer Science, Stanford)

I. Summary of research program

A. Technical goals

The goals of the protein structure modeling project are to 1) identify critical tasks in protein structure elucidation which may benefit by the application of AI problem-solving techniques, and 2) design and implement programs to perform those tasks. We have identified two principal areas which have both practical and theoretical interest to both protein crystallographers and computer scientists working in AI. The first is the problem of interpreting a three-dimensional electron density map. The second is the problem of determining a plausible structure in the absence of phase information normally inferred from experimental isomorphous replacement data. Current emphasis is on the implementation of a program for interpreting electron density (e.d.) maps.

B. Medical relevance and collaboration

The biomedical relevance of protein crystallography has been well stated in an excellent textbook on the subject (Blundell & Johnson, Protein Crystallography, Academic Press, 1976):

> "Protein Crystallography is the application of the techniques of X-ray diffraction ... to crystals of one of the most important classes of biological molecules, the proteins. ... It is known that the diverse biological functions of these complex molecules are determined by and are dependent upon their three-dimensional structure and upon the ability of these structures to respond to other molecules by changes in shape. At the present time X-ray analysis of protein crystals forms the only method by which detailed structural information (in terms of the spatial coordinates of the atoms) may be obtained. The results of these analyses have provided firm structural evidence which, together with biochemical and chemical studies, immediately suggests proposals concerning the molecular basis of biological activity."

The project is a collaboration of computer scientists at Stanford University and crystallographers at the University of California at San Diego (under the direction of Prof. Joseph Kraut) and at Oak Ridge National Laboratories (Dr. Carroll Johnson). Our principal collaborator at UCSD is Dr. Stephan Freer.

C. Progress summary

During the past year we have continued with the design and implementation of a system of programs for interpreting three-dimensional e.d. maps. Progress has been made by attacking the problem from two directions: working upward from the primary data (i.e., the array of e.d. values) to higher level symbolic abstractions, and working downward from the given amino acid sequence and other experimental information to generate candidate structures which can then be confirmed by the abstracted data.

Research which emphasized the "bottom-up" approach yielded results in several areas:

1) A new skeletonization technique, based on Greer's work, was implemented, making it considerably easier to generate an optimum skeletal representation of the map. Experience in using this method on several different e.d. maps has led to heuristics for selecting grid size, minimum density threshold, contour levels, etc. Moreover, some additional post-processing of the skeleton has been implemented to indicate connectivity, to produce a list of "model atoms" for map refinement (see below) and to provide input for graphical display.

2) A new technique for improving the quality of medium resolution (about 2.5 Angstroms) e.d. maps was formulated and tested on simple structures and on a real protein currently under investigation at UCSD. The technique utilizes the quasi-atomic structure produced by skeletonizing the map. The grid points in the map which are not removed by the skeletonization process are treated as zeroth-order model atoms. A difference map is then computed, using the difference between observed and calculated structure amplitudes, and calculated phases, for the Fourier coefficients. The difference map is then used to move the model atoms toward regions of higher density. The process is repeated until the shifts in atomic coordinates become insignificant. The standard crystallographic R value is used as one measure of quality of the map. Another measure is a subjective one of displaying the map and inspecting it in regions where the protein structure has been partially determined. This technique is currently being used as one step in the determination of the structure of cytochrome c peroxidase. The method is similar to one formulated by Agarwal in its use of a set of model atoms, whose positions are progressively refined. The use of a difference map is based on earlier methods of refinement developed by Freer et al.

3) We investigated the utility of graphics systems as a tool for acquiring model building heuristics. One system is a grey-scale display used in image understanding research at the Stanford Artificial Intelligence Laboratory. A second system is the Picture System used by Prof. Langridge's group at the UCSF Medical Center. We found the former system to be difficult to use as an interpretation aid, due to the necessity of visually synthesizing two-dimensional slices of the e.d. map. We attempted to incorporate depth into the display by making the grey-scale intensity a function of both electron density and depth, but the technique wasn't helpful. Other schemes were considered, but not implemented, due to the inordinate amount of programming effort that would be required. The Picture System at the UCSF Medical Center is similar in many respects to the system used by our UCSD collaborators, and its proximity to Stanford makes it attractive. We found the system to be a potentially valuable aid in comparing the skeletal representation of the e.d. map with the more

conventional contour map representation, as well as comparing the skeleton with known protein structures. At present, however, there is no working software for displaying a contour map and a skeletal model simultaneously, but an effort is underway to put this capability into the system.

The development of our "top-down" system (named CRYSALIS) for inferring the molecular structure from the amino acid sequence, symbolic abstractions of the e.d. map and stereochemical knowledge, is continuing. Recent additions to the system include rules for identifying main chains, side chains and bridge segments, using knowledge of expected topological properties of the skeleton and peak distributions. A new report, which focusses on the design of the CRYSALIS system, will be out shortly. The purpose of that report is to summarize the current state of the system, and to critically review it with respect to its design specifications. Although the system does make inferences from the data about some structural features of the model, it has been difficult to extend the power of the system beyond its present level. The design review was primarily motivated by a desire to see which features of the system are worth preserving and which ones need redesigning in the next version of CRYSALIS.

References:

Agarwal, R. C., and Isaacs, N. W., "Method for obtaining a high resolution protein map starting from a low resolution map", Proc. Natl. Acad. Sci. USA, Vol. 74, pp 2835-2839 (1977).

Freer, S. T., Alden, R. A., Carter, C. W. and Kraut, J., "Crystallographic Structure Refinement of Chromatium High Potential Iron Protein at Two Angstroms Resolution", J. Biol. Chem., Vol. 250,46 (1974).

Greer, J., "Three-dimensional Pattern Recognition: An Approach to Automated Interpretation of Electron Density Maps of Proteins", J. Mol. Biol., Vol. 82, pp. 279-301 (1974).

D. List of Publications

1) Robert S. Engelmore and H. Penny Nii, "A Knowledge-Based System for the Interpretation of Protein X-Ray Crystallographic Data," Heuristic Programming Project Memo HPP-77-2, January, 1977. (Alternate identification: STAN-CS-77-589)

2) E.A. Feigenbaum, R.S. Engelmore, C.K. Johnson, "A Correlation Between Crystallographic Computing and Artificial Intelligence," in Acta Crystallographica, A33:13, (1977). (Alternate identification: HPP-77-25)

E. Funding status

Grant title: The Automation of Scientific Inference: Heuristic Computing Applied to Protein Crystallography

Principal Investigator: Prof. Edward A. Feigenbaum

Funding Agency: National Science Foundation

Grant identification number: MCS 74-23461-A01

Term of award: May 1, 1977 through April 30, 1979

Amount of award: $150,200 (including indirect costs)

II. Interaction with the SUMEX-AIM resource

A. Collaborations

The protein structure modeling project has been a collaborative effort since its inception, involving co-workers at Stanford and UCSD (and, more recently, at Oak Ridge). The SUMEX facility has provided a focus for the communication of knowledge, programs and data. Without the special facilities provided by SUMEX the research would be seriously impeded. Computer networking has been especially effective in facilitating the transfer of information. For example, the more traditional computational analyses of the UCSD crystallographic data are made at the CDC 7600 facility at Berkeley. As the processed data, specifically the e.d maps and their Fourier transforms, become available, they are transferred to SUMEX via the FTP facility of the ARPA net, with a minimum of fuss. (Unfortunately, other methods of data transfer are often necessary as well -- see below.) Programs developed at SUMEX, or transferred to SUMEX from other laboratories, are shared directly among the collaborators. Indeed, with some of the programs which have originated at UCSD and elsewhere, our off-campus collaborators frequently find it easier to use the SUMEX versions because of the interactive computing environment and ease of access. Advice, progress reports, new ideas, general information, etc. are communicated via the message and/or bulletin board facilities.

B. Interaction with other SUMEX-AIM projects

Our interactions with other SUMEX-AIM projects have been mostly in the form of personal contacts. We have strong ties to the DENDRAL, Meta-DENDRAL and MOLGEN projects and keep abreast of research in those areas on a regular basis through informal discussions. The SUMEX-AIM workshops provide an excellent opportunity to survey all the projects in the community. Common research themes, e.g. knowledge-based systems, as well as alternate problem-solving methodologies were particularly valuable to share.

C. Critique of Resource services

On the whole the services provided by SUMEX have been excellent, considering the large demand on its resources. With the important exceptions of high peaks in the weekday prime-time load average, the ratio of CPU time to total wait time during program execution is usually acceptable. The facility provides a wide spectrum of computing services which are genuinely useful to our project -- message handling, file management, Interlisp, Fortran and text editors come immediately to mind. Moreover, the staff, particularly the operators, are to be commended for their willingness to help solve special problems (e.g., reading tapes) or providing extra service (e.g., and immediate retrieval of an archived file). Such cooperative behavior is rare in computer centers.

The most serious deficiency, from our point of view, is the lack of a file transfer facility between SUMEX and the computing system in the UCSD Chemistry Department. Our day-to-day collaboration with Dr. Freer at UCSD would be greatly enhanced by a reasonably fast (even 1200 baud would suffice) channel for transmitting proposed protein models, generated at SUMEX, to the Picture System at UCSD.

III. Use of SUMEX during the follow-on grant period (8/78 - 7/83)

A. Long-range goals

Our current research grant extends through April, 1979. During that time we intend to bring the structure modeling system to a level of performance that permits reliable qualitative interpretation of high resolution e.d. maps, derived from real data and a correct amino acid sequence. We also plan to exploit the flexibility of the rule-based control structure to permit investigation of alternate problem-solving strategies and modes of explanation of the program's reasoning steps. Beyond the next two years, emphasis will be placed on expanding and generalizing the system to relax the constraints of resolution and accuracy in the input data.

B. Justification for continued use of SUMEX

The biomedical relevance of the protein structure modeling project, coupled with the need for building a computational system with a significant component of symbolic inference, qualifies the project as an AIM-relevant endeavor. SUMEX provides an excellent computing environment for creating and debugging programs (in a variety of languages), for sharing and distributing information among geographically dispersed co-workers, and for keeping up with current research in other AIM areas. Our project is clearly too small to justify an independent computing facility, and other large computer centers that are conveniently accessible do not fulfill our requisites. Consequently SUMEX has been and hopefully will continue to be an integral research tool in this project.

4.3 PILOT AIM PROJECTS

The following are descriptions of the informal pilot projects currently using the AIM portion of the SUMEX-AIM resource pending funding, and full review and authorization.

4.3.1 COMMUNICATION ENHANCEMENT PROJECT

Communication Enhancement Project

John B. Eulenberg, Ph.D. and Carl V. Page, Ph.D.
Department of Computer Science
Michigan State University

I) Summary of research program.

A) Technical goals.

The major goal of this research is the design of intelligent speech prostheses for persons who experience severe communication handicaps. Essential subgoals are:

(1) Design of input devices which can be used by persons whose movement is greatly restricted.

(2) Development of software for text-to-speech production.

(3) Research in knowledge representations for syntax and semantics of spoken English in restricted real world domains.

(4) Development of micro-computer based portable speech prostheses.

B) Medical Relevance and Collaboration.

Members of our group are in touch with Dr. Kenneth Colby and his group at UCLA who are working on similar problems for a domain of people who have aphasia.

The need for such technology in the medical area is very great. Millions of people around the world lead isolated existences unable to communicate because of stroke, traumatic brain injury, cerebral palsy or other causes. The availability of inexpensive micro-processors and voice synthesizers allows development of complex experimental systems to study human communication. The knowledge gained from these experimental systems should lead in a few years to prototypes of very low cost which will permit many people to engage in the vital acts of communication required for a "normal" life in human society.

We have organized institutes to bring together the many professionals who have an interest in this area. Together with the Tufts New England Medical Center, the TRACE Center for Research and Development for the Severely Communicatively impaired of U. of Wisconsin, and the Children's Hospital at Stanford (Maurice LeBlanc). We have begun the first newsletter for dissemination in this area. Called "Communication Outlook", the first issue will be published in April, 1978. Subscribers and contributors to the Newsletter come from a vide variety of disciplines and from many countries. John B. Eulenberg helped to organize the first Federal workshop for governmental agencies who have some interest in funding work in these areas. Represented were the Bureau of Education for the Handicapped, The Veterans Administration, The Civil Service Commission, NIH, NSF, and others. We have also been in touch with United Cerebral Palsy associations at the state and national levels. Much of our effort has been in educating those medical, educational, and governmental communities with an interest in this area on the available technology since most of them are not accustomed to funding the development of high-technology systems.

C) Progress summary.

Although some facets of the research have been underway at MSU for several years, we have been using SUMEX-AIM for only about a year at this time, having received our password in March, 1977. During the last year, we have:

1) Designed and built hardware and software allowing us to transmit files to SUMEX from our Nova 2/10 at 300 baud.

2) Organized a research team of 4 students possessing background in artificial intelligence lead by Dr. Carl V. Page to start a semantics-speech generator. This group had a very primitive prototype (written in running in June, 1977. The system uses statistical, grammatical and semantic information to generate sentences by anticipation. We are organizing a similar group again this month (we have a seasonal supply AI students) to expand the semantics.

3) Converted a large program (Orthophone) for English text to speech synthesizer codes to SAIL from Algol.

4) Obtained local support for terminals and tie-lines to use the SUMEX-AIM facility. We requested these in our original proposal but were not granted them. At present, the lack of a dedicated tie-line from East Lansing to Tymshare in Ann Arbor or Detroit is a problem for us during 0600 to 0900 PST.

5) During the past year, Dr. Reid of our project refined a wheel-chair portable personal communication system for a 10 year old boy who has cerebral palsy. It is micro-computer based and can accept inputs via an adaptive switch from a series of menus displayed on a TV screen, via Morse code, or by a keyboard. Its outputs can be TV display, hard copy, spoken English, Morse code, or musical sounds. As the memory available for small systems will soon be substantial, we will need to specify the content and connection of the choice menus using the knowledge gained in our SUMEX-AIM project.

6) A Doctoral thesis in the association of knowledge sources (letter and word frequencies, syntactics, semantics, pragmatics and belief systems) for the generation of speech has been started by one of our students, Mr. James Soddy (Supervised by Carl V. Page). Mr. Soddy will use the SUMEX-AIM system during Summer 1978 to program some examples for his thesis, as a means of obtaining current information from the AI community, and to communicate with Dr. Page who will be working in Sunnyvale, Calif. for the summer.

7) We have built and tested a myoelectric interface and used it (together with a miniature FM transmitter) for input of changing muscle potentials into a computer. There is reason to believe that this means of input may provide a higher bit rate than other known means for those people who possess severe cerebral palsy.

8) We have developed software for teaching basic educational concepts to severely impaired persons. For example we have developed a "talking" system for drilling students in Bliss symbolics. Another system we have developed teaches spelling using a voice synthesizer and TV screen.

D) Up-to-date list of publications. (1976 to date)

By John B. Eulenberg

"Technical Systems Development, Headin", Interim Report, April, 1976, Experimental Applications of Two-Way Cable Delivery, NSF Grant No. APR 75-14286.

"Interactive New Hired Information Access System with Both Voice and Hard Copy Output: User's Guide to NHQUERRY", April 11, 1976 (With Steven Kludt and Jerome Jackson (Artificial Language Laboratory Report AEB 041176))

"Language Individualization in a Computer-Based Speech Prosthesis System", National Computer Conference, New York, June 9, 1976.

"Individualization in a Speech Prosthesis System", Proceedings of 1976 Conference on Systems and Devices for the Disabled, June 10, 1976.

"The LEAF Language", Interim Report, September, 1976, NSF Grant No. APR 75-14286.

"Microprocessor-Based Artificial Language for Communication Prostheses", with M. R. Rahimi, Proc. of the National Electronics Conference, Vol. XXXI, October, 1977.

"A programmable Multi-Channel Modem Output Switch", September 22, 1976, with Joseph C. Gehman and Juha Koljonen (Artificial Language Laboratory Report AEB 092276)

"SMPTE Time Code Interface and Computer-Controlled Video Switcher", with Michael Gorbutt and Dennis Phillips, Interim Report, March, 1977 NSF Grant APR 75-14286.

"Representation of Language Space in Speech Prostheses", with R. Reid and M. Rahimi, Proc. of Fourth Annual Conference on Systems and Devices for the Disabled, June, 1977.

"Administration and Management of a Computer-Based Communication Enhancement Program", with M. R. Rahimi and L. Neiswander, Proc. of Amer. Acad. for Cerebral Palsy and Developmental Medicine, October, 1977. "When [-VOICE] becomes [+VOICE]- The Phonological Competence of People Who Cannot Speak", with Carol Myers Scotton, Proceedings of the Annual Confer. of the Linguistic Soc. of America, December, 1977.

By Carl V. Page:

"Heuristics for Signature Table Analysis as a Pattern Recognition Technique", IEEE Transactions on Systems, Man and Cybernetics,Vol. SMC-7, No. 2, February 1977.

"Discriminant Grammars, an Alternative to Parsing". with Alan Filipski, Proceedings of the IEEE Workshop on Picture Processing, Computer Graphics, and Pattern Recognition, April 22, 1977.

"Pattern Recognition and Data structures". Chapter in "Data Structures in Computer Graphics and Pattern Recognition" Edited by Allen Klinger, Academic Press, 1977. "A Survey of Artificial Intelligence in Computer-Aided Instruction", with Alice Gable (Submitted to the International Journal of Man-Machine Systems, March, 1978)

E) Funding Status.

1) Current funding. Wayne County (Detroit)

Wayne Intermediate School District	$75,816	(Third year)
Northville Public School District	$41,333	(Third year)
Jackson Intermediate School District	$26,500	(Second year)
Ingham Intermediate School District	$23,700	(First Year)
Michigan State University Division of Engineering Research	$64,500	(for each of two years).
Grand Rapids Public Schools	$2,100	
Vandervoot Foundation	$5,000	

Some of this money has been used to purchase equipment which is the property of WCISD for use in a demonstration classroom in an elementary school. Commitments in the grants have prevented us from using very much of these funds to support long range goals such as those communicated to SUMEX-AIM. However, the special communication devices, student and other research facilities provide the critical mass which will allow us to do the work that we have proposed. The main value of SUMEX-AIM to us is to allow experimentation with AI technology in order to develop the theory to develop intelligent speech prostheses.

2) Pending applications and renewals.

Oakland County Intermediate School District (Application being considered after negotiation)	$100,000.
Genessee County Intermediate School District	$100,000.
Tuscola County Intermediate School District	$20,000.
Livingston County Intermediate School District	$50,000.

As one can see from this list of sources, there is a lot of interest in this area from agencies which are not experienced in funding high-technology and research.

II) Interactions with the SUMEX-AIM resource.

A) Examples of medical collaboration and medical use of programs via SUMEX.

The faculty in the MSU College of Human Medicine who teach medical decision making were shown a demonstration of the SUMEX system, MYCIN and PARRY. We have agreed to present a demonstration of PUFF to Dr. Clyde Flory, an allergist who is the most knowledgeable person in our area in pulmonary studies. We intend to explore the possibility of a table-driven program for the treatment of allergy with Dr. Flory. If we decide to undertake the development of this, we will send another proposal to SUMEX-AIM. A member of our Medical School faculty, Dr. Richard Ropple, an expert on myoelectronics, is a member of of our research group.

The Dean of our College of Human medicine visited our laboratory in April, 1977 and we have been asked to apply for inclusion in the University's Clinical Center as part of the Rehabilitation Medicine Program.

B) Examples of sharing, contacts, and cross-fertilization with other SUMEX-AIM projects.

1. We have met with Dr. Kenneth Colby on many occasions including the SUMEX-AIM workshop in June, 1976. Our work in many ways complements his and we have had several worthwhile interchanges of information. We are converting our major software for speech generation and adaptive inputs to the SUMEX-AIM system in part so that they can be used by Dr. Colby and his group.

2. Mr. Douglas Appelt, a doctoral student at SU-AI was our principal systems programmer last summer. He is currently doing research in the same area as ours with DR. Gary Hendrix of SRI. We have used his knowledge of your system (via the message sending routines) to assist us in starting our project.

C) Critique of resource services.

Our use of SUMEX-AIM has been seasonal with most programs run during Spring Term. We have used your system for work that could not be done conveniently on our computers. We have been pleased with the system and find it as easy to use as one that is close to us geographically. Dr. Page will be working in the Bay Area this summer and plans to visit your facility as well as use it to keep in touch with the work at East Lansing.

III) Follow-on SUMEX grant period (8/78-7/83).

A) Long-range user project goals and plans.

We want to do fundamental research in artificial intelligence in the context of the generation of speech from very minimal amounts of input. This problem seems closely related to the understanding of speech. It seems that the methods of representation of knowledge used for speech or vision understanding can be used in a natural way for fluent generation of speech. Our area seems almost unique in AI in that it is socially desirable (without question). Even relatively primitive systems can improve the quality of life for hundreds of thousands of people.

Major long range goals are:

1) To do research in transposing the vocal tract to another region of the body in which an individual has suitable myoelectric control for the generation of speech.

2) To define a suitable system of semantics and to encode world knowledge in that system that would be useful for the generation of speech fluently.

3) To discover primitive operations on semantics which allow new and appropriate combinations of speech to be generated. (Using other sources of knowledge.)

4) To develop means for a severely handicapped individual to program and personalize his or her own speech and environmental control system.

5) To study means of using speech output to aid blind persons both through experiments with simplified text to speech devices and through means of teaching blind persons to write in cursive.

6) To study the effect of communication aids technology on the psychological assessment of individuals previously thought to be retarded and to study the consequences of this technology for the educational system.

7) To improve the prosodic qualities of generated speech, using its semantic aspects.

8) To design portable speech prostheses which allow maximum use of state of the art knowledge in speech generation.

9) To develop an experimental base for studying how the concepts which are articulated in speech are manipulated by individuals at differing states of mental organization

10) To study the potential for speech generation systems as a means of stimulating autistic children.

11) To develop voice recognition systems which will aid individuals with limited speech to develop their full potential.

(We don't expect to finish all of these by 1983.)

B) Justification for continued use of SUMEX by your project.

1) We need to use many sources of knowledge represented in computers to do our work, similar to many SUMEX users.

2) We know kindred spirits in the AI community who possess goals similar to our long range goals.

3) We have substantial hardware and software expertise which we are willing to share.

4) The payoff to society of our research in terms of the improved quality of millions of human lives seems great.

5) This area does not have a traditional means of support for research separate from development which makes your support vital for our long range goals. 6) Our area is very interdisciplinary and the communication aspects of SUMEX-AIM will be increasing valuable to us.

C) Plans For Other Computational Resources.

We have available to us three mini-computers: a Nova 3/10, a PDP-11/34 and a PDP-11/45 as well as the CDC 6500 and CDC 6400 of our central computer facility. Our demonstration classroom in a Detroit suburb will open soon making a Nova available to students who experience severe communication and motoric handicaps. None of our small systems possess AI software. We hope to develop prototype systems on SUMEX-AIM within the next year the can be used in our demonstration classroom. We believe that the injection of quite small amounts of AI technology into the speech generation systems can produce significant improvements in the communication and educational processes. We will be in a position to measure the effectiveness of the AI tools which we try. If we can make a case on a cost effectiveness basis, there are sources of support to acquire appropriate hardware to service a larger group of students and also support our research. At this point we do not feel the need to try to acquire a PDP-10 or other machine suited to AI research because although important to us, the number of machine cycles we require is relatively small.

D) Recommendations for Future Community and Resource Development.

1) We would find it helpful if there were means within SUMEX-AIM to assist in the transfer of prototypes written for your system to various common minicomputers.

2) We have available software to assist handicapped individuals in using computer systems. We would be happy to facilitate the use of SUMEX-AIM resources by a blind programmer by modifying appropriate software.

4.3.2 COMPUTERIZED PSYCHOPHARMACOLOGY ADVISOR

A Computerized Psychopharmacology Advisor

Jon F. Heiser, M.D.
Dept. of Psychiatry and Human Behavior
University of California at Irvine

I. Summary Research Program

A. Technical Goals

We propose to develop a computer-based automated system for education and consultation in clinical psychopharmacology. Our technical goals are envisioned in three phases:

- To develop a model of expert teaching, consulting and decision-making in clinical psychopharmacology.
- To implement this model on a computer system which responds in real time and communicates in natural language.
- To evaluate the performance of the system as a teaching and consulting aid.

B. Medical Relevance and Collaboration.

1. Medical Relevance.

For many years, it has been recognized that potent psychopharmacological agents are frequently used in an unsystematic manner. There are at least 50 discrete syndromes currently identified in clinical psychiatry which have unique hierarchies of plausible pharmacological treatments. Each therapeutic regimen in each hierarchy may involve several classes of drugs which can be preferentially ranked. A particular member of a class of drugs may occasionally be recommended on the basis of a patient's medical history, family history, response to previous treatments, current physical status, or current mental status. In addition, each treatment program has its own set of potential side effects, adverse reactions and drug-drug, drug-host, drug-age, drug-gender, drug-state-of-health, and drug-other treatment interactions.

Conventional sources of information for education or verification (books, journals, lectures, and seminars) are seldom quickly accessible or specifically pertinent. A traditional alternative is to consult a specialist. In addition to availability, reliability and validity, a good consultant has the ability to understand questions in their proper context and sequence, to give advice which can be explained or documented as needed, and to provide follow-up consultations which incorporate new information from clinical developments or additional expertise.

Our research on the Clinical Psychopharmacology Advisor is directed towards implementing all of the characteristics of a good consultant, which have only been outlined above, in a functional computer program. To our knowledge, no other computer program currently available or under development fulfills all of these goals in clinical psychopharmacology.

2. Collaboration.

a. Principal Investigator: Jon F. Heiser, M.D., Associate Clinical Professor, Department of Psychiatry and Human Behavior, University of California, Irvine.

b. Co-principal Investigator: Ruven E. Brooks, Ph.D., Assistant Professor, Department of Information and Computer Science (ICS) and Department of Psychiatry and Human Behavior, University of California, Irvine.

c. Resident Physicians, Department of Psychiatry and Human Behavior, University of California, Irvine:

Bronco R. Radisavljevic, M.D.	(March 1977)
Steven J. Smith, M.D.	(June 1977)
Frank S. Floca, M.D.	(January-March, 1978)
Neal R. Cutler, M.D.	(April-June, 1978)

d. Research Associate: Joan P. Ballard, Ph.D.

e. Medical Students, California College of Medicine, University of California, Irvine:

Clifford Risk	(October-December, 1975)
Dana W. Ludwig	(October-December, 1975)
Sue A. Clear	(May-September, 1976)
Neil R. Shocket	(April 1977-present)

f. Pharmacists, University of California Irvine Medical Center:

Pierre J. Menard, Pharm. D.	(January-June 1977)
Michael S. Toole, Pharm. D.	(September 1976-present)

g. Undergraduate Computer Science Majors, Department of Information and Computer Science, University of California, Irvine.

Darryl Hansen	(March-June, 1977)
Thomas E. Holthus	(March 1977-present)

h. Local Advisory Panel, Department of Psychiatry and Human Behavior, University of California, Irvine:

Louis A. Gottschalk, M.D.
Edward Kaufman, M.D.
John Kramer, M.D.
David Shore, M.D.

i. National Advisory Panel:

Ross J. Baldessarini, M.D.
Neuropharmacology Laboratory
McLean Hospital
Brookline, Massachusetts 02168

John M. Davis, M.D.
Illinois State Psychiatric Institute
1601 West Taylor Street
Chicago, Illinois 60612

Max Fink, M.D.
Department of Psychiatry
State University of New York at Stony Brook
Stony Brook, Long Island, New York 11794

Frederick Goodwin, M.D. (Tentative)
National Institute of Mental Health
9000 Rockville Pike, Building 10, Room 3S205
Bethesda, Maryland 20014

John H. Greist, M.D.
Department of Psychiatry
University of Wisconsin
Madison, Wisconsin 53706

Leo E. Hollister, M.D.
Departments of Medicine and Psychiatry, Stanford University
Veterans Administration Hospital
Palo Alto, California 94302

James W. Jefferson, M.D.
Department of Psychiatry
University of Wisconsin
Madison, Wisconsin 53706

Donald F. Klein, M.D.
New York Psychiatric Institute
722 West 168th Street
New York, New York 10032

George M. Simpson, M.D.
Psychopharmacology Unit, University of Southern California
Metropolitan State Hospital
Norwalk, California 90650

Robert L. Spitzer, M.D.
New York Psychiatric Institute
722 West 168th Street
New York, New York 10032

Zebulon C. Taintor, M.D.
Rockland Research Institute
Rockland State Hospital
Orangeburg, New York 10962

C. Progress Summary.

After carefully reviewing alternative contexts in which to generate the Psychopharmacology Advisor, we selected the EMYCIN software. MYCIN is a computer program which functions as a consultant in the diagnosis and treatment of infectious diseases. EMYCIN is a version of the MYCIN system with all knowledge and reference to infectious disease removed but with all features for diagnosis, treatment recommendation, explanation and knowledge acquisition retained. Our choice of EMYCIN was determined primarily by the availability of the EMYCIN software (including a significant amount of professional consultation, collaboration and system maintenance supplied by the MYCIN staff), the suitability of EMYCIN to most of our initial design considerations, and the desire of the MYCIN staff to test the EMYCIN software in a different clinical domain. EMYCIN has thus become our working model of expert teaching, consulting and decision-making in clinical psychopharmacology.

Our initial goal has been to develop a small but fully functioning Clinical Psychopharmacology Advisor. Approximately 250 rules, utilizing about 120 clinical parameters, have been developed and are used to diagnose and recommend therapy. The system, informally called HEADMED, currently has sound knowledge about the differential diagnosis of the major affective disorders, schizophrenia and the general category of organic brain disorders. The Psychopharmacology Advisor has perfunctory information concerning paranoid disorders and personality disorders. The system has a rudimentary knowledge of the Minnesota Multiphasic Personality Inventory (MMPI). HEADMED presently has skeletal knowledge about neuroses, behavior disorders and substance abuse and about organic brain disorders, regarding both type of brain disorder (e.g. delirium or dementia) and cause of brain disorders (e.g. intoxication or trauma. the program knows nothing about child psychiatry, sexual disorders and other psychiatric conditions.

The HEADMED software currently has the capability of recommending a drug treatment, if indicated, and of cautioning about potentially harmful interactions with a compromised host and with other chemical substances. The system also can print out advice concerning dosage and duration of therapy, pharmacokinetics, and warnings about common side effects and possible adverse reactions. We are beginning to develop data structures which can utilize this knowledge to compute diagnostic formulations and therapeutic plans which are highly specific to the unique properties and circumstances of a particular patient.

Work in progress is also concerned with improving the data input characteristics of the system. For example, it is common to describe a patient's thought pattern as "loose". A good tutor or consultant would immediately ask how loose: at the "paragraph" level (hardly a disorder as anyone composing a manuscript can certify), at the sentence level, the clause level, or within a word (possibly producing neologisms)? Furthermore, an experienced teacher might ask a novice student to discriminate between a loose association which is creative, a loose association which is humorous, and a loose association which is pathological.

As an initial step towards describing clinical phenomena reliably and validly, we are greatly expanding HEADMED's synonym list by interviewing various clinicians, abstracting a large number of psychiatric hospital charts, and studying results of computer consultations with HEADMED both for unanticipated responses and accepted but misinterpreted responses. Within the system, we are making extensive use of EMYCIN features called SPECIAL and EXTRASPECIAL to better anticipate how various users might describe patients and to quiz them immediately when their responses are ambiguous or unanticipated.

Another project consists of continuing to expand and modify the rules for diagnosis and treatment recommendation. Currently we are experimenting with rules which have more elaborate descriptions of the clinical disorders in their premise or left-hand side. We are just beginning an extensive elaboration of the pharmacological content of the system.

Work has also begun on modifying the EMYCIN therapy module, our first attempt at actually changing the basic form of the software. Presently, EMYCIN selects therapies by multiplying the certainty factor (CF) of each diagnosis by the CF of every treatment for that diagnosis. In most cases of psychiatric interest, there are several probable diagnoses and relatively few classes of psychopharmacological medications. It is common, therefore, for the same treatment to be recommended several times. Using the standard EMYCIN therapy algorithm, too many treatments acquire high CF's. Not uncommonly, therapeutic recommendations for subordinate diagnoses attain higher CF's than a less common but more specific and presumably better treatment recommendation for the primary diagnosis.

We have modified the EMYCIN software so that only the top diagnosis is used in choosing a therapy. Later, we plan to give the user various options such as treatment for top diagnosis only, treatment for any other diagnosis alone, or "broad-spectrum" treatment for any combination of the diagnoses (listed or unlisted) in any hierarchical order (for example, "I don't like diagnosis 2 at all, so throw that out. Now what would be your recommendation if the remaining diagnoses were inverted in rank? Suppose that the primary diagnosis were Mania, which you didn't even consider; how would that affect your recommendations?")

Demonstrations of the Psychopharmacology Advisor have been given to numerous groups and individuals on an informal basis. The following formal demonstrations of HEADMED were given in the past year:

1. VI World Congress of Psychiatry, Honolulu, Hawaii, 29 Augsut 1977.

2. Department of Computer Science, University of Hawaii, Honolulu, Hawaii, 01 September 1977.

3. Department of Psychiatry, University of Wisconsin, 21 September 1977.

4. Richard F. White, M.D., Director, Squibb Professional Services, 11 October 1977.

5. George M. Simpson, M.D., Professor of Psychiatry, University of Southern California, Metropolitan State Hospital, Norwalk, California, 03 November 1977.

6. Rockland Research Institute, Orangeburg, New York, 05 January 1978.

7. New York Psychiatric Institute, New York, New York, 06 January 1978.

8. NIMH Site Visiting Team, Extramural Psychopharmacology Research Division, 23 January 1978.

9. Department of Psychiatry, University of Texas Medical Branch, Galveston, Texas, 02 February 1978.

10. American Psychiatric Association Annual Meeting, Atlanta Georgia, 10 May 1978.

Many useful suggestions and critical insights are generated by these demonstrations.

D. List of Relevant Publications.

1. Heiser, J.F., Colby, K.M., Faught, W.S. and Parkison, R.C. Can Psychiatrists Distinguish a Computer Simulation of Paranoia from the Real Thing? Submitted for publication.

2. Heiser, J.F., Brooks, R.E. and Ballard, J.P. Artificial Intelligence in Psychopharmacology. Abstracts - VI World Congress of Psychiatry, Honolulu, Hawaii, 28 August - 03 September 1977, page 135.

3. Heiser, J.F., Brooks, R.E., and Ballard, J.P. A Computerized Psychopharmacology Advisor. Scientific Proceedings in Summary Form, The 131st Annual Meeting of the American Psychiatric Association, Atlanta, Georgia, May 1978 (in press).

4, Heiser, J.F., Brooks, R.E., and Ballard, J.P. Progress Report: A Computerized Psychopharmacology Advisor. Proceedings of the 11th Collegium Internationale Neuro-Psychopharmacologicum, Vienna, Austria, July 1978 (in press).

5. Verbal Reports:

   a. Heiser, J.F. Computer-Aided Diagnosis of Psychiatric Patients. Presented to the Research Meeting, School of Engineering, University of California, Irvine, 07 October 1976.

   b. Brooks, R.E. and Heiser, J.F. An application of Artificial Intelligence to Psychiatry. Presented to:

      Indian Institute of Technology, Madras, India, 28 September 1976, and

      Madras Christian College, Madras, India, 03 October 1976.

E. Funding Support Status.

1. The Principal Investigator, Co-Principal Investigator, Resident Physicians and Pharmacists have all been full-time employees of the University of California, Irvine or the Veterans Administration, and have participated in this research as part of their assigned duties or in their spare time.

2. Additional support in the form of Office and Laboratory Space, Clerical Assistance, Peripheral Data Processing Equipment, Supplies and Expenses for Traveling to Professional Meetings has also been provided by the University of California, Irvine.

3. The Medical Students have worked on this project either during elective periods for academic credit or during free time with support from NIMH Undergraduate Training Program Grants:

   Title: A Computerized Psychopharmacology Advisor
   Principal Investigator: Jon F. Heiser, M.D.
   Funding Agency: National Institute of Mental Health (NIMH)
   Undergraduate Training Program
   Total Award: $2400
   (2 $600 awards to Sue Arrigo Clear and 2 $600 awards to Neil R. Shocket)
   Date: 1976-present

4. Undergraduate Student Darryl Hansen worked on this project during UCI ICS Department Course 199, Independent Study. Undergraduate Student Thomas E. Holthus worked on this project during UCI ICS Department Course 199, Artificial Intelligence Project. Mr. Holthus also spent the Summer Quarter 1977 working on the project with support from the National Science Foundation (NSF) through an Undergraduate Research Training Grant to the Department of Information and Computer Science.

5. Since October 1977, Mr. Holthus has been employed half-time by the University of California, Irvine as a Research Technician in the Department of Psychiatry and Human Behavior. Mr. Holthus is assigned to work on this project and is paid $3.67 per hour.

6. A modest amount of additional support has been obtained from:

   Title: A Computerized Psychopharmacology Advisor
   Principal Investigator: Jon F. Heiser, M.D.
   Funding Agency: Anne R. Issler Endowment Fund
   Department of Psychiatry and Human Behavior
   University of California, Irvine
   Total Award: $552.50
   Dates: Awarded 06 January 1978 for an indefinite time period

7. A grant application was submitted to the National Institute of Mental Health on 01 November 1977. A grant application was submitted to the Veterans Administration on 01 January 1978. Copies of these grant proposals are available from the Principal Investigator.

8. The Director of Professional Services, E.R. Squibb and Sons Pharmaceutical Company, has offered to support Professional Collaboration through Squibb's panel of distinguished consultants.

II. Interactions with the SUMEX-AIM Resource.

A. Collaborations and Medical Use of Programs via SUMEX.

1. The MYCIN group has collaborated with our group since work on the Psychopharmacology Advisor began. The MYCIN group supplies invaluable software support to the EMYCIN program. Our group has participated in writing documentation of the EMYCIN software which presumably will be useful to all EMYCIN users. In addition to discussions of various issues by LINKs, MESSAGEs, phone calls and attendance at workshops, Dr. Heiser and Dr. Brooks have made several trips to the MYCIN headquarters, and Mr. Fagan of the MYCIN group has visited the HEADMED team.

B. Sharing and Interactions with Other SUMEX-AIM Projects.

1. Collaboration with Kenneth Mark Colby, M.D. and members of the Higher Mental Functions Project begun last year has continued in the form of writing and submitting for publication a paper reporting a "Turing Test" which was performed on-line on SUMEX, with the psychiatrist-judges located at the University of California, Irvine, the patient-person at the University of California, Los Angeles (UCLA) and PARRY at SUMEX. Prepublication copies of this paper (see I.D.1. above: Heiser et al. Can Psychiatrists Distinguish a Computer Simulation of Paranoia from the Real Thing?) are available upon request. In addition, demonstrations of the PARRY and DOCTOR programs have been given on-line, using SUMEX, to various groups of mental health professionals, computer scientists and other qualified and interested individuals.

2. The Principal Investigator visited Dr. Eulenberg and Dr. Page of the Communication Enhancement Project at Michigan State University on 23 September 1977.

3. As noted last year, the Principal Investigator first became acquainted with the SUMEX-AIM resource and initiated his collaboration with the MYCIN project by attending the first SUMEX-AIM Workshop at Rutgers University in June 1975. Dr. Heiser has continued to benefit from these workshops.

C. Critique of Resource Management.

In general, we find the SUMEX resource a hospitable environment. Despite periods of heavy load, we find that sufficient resources are available when we need them. The operating system and associated utility subsystems are flexible and powerful, and the mail system is uniquely valuable. Finally, we consider the practice of keeping system documentation on-line to be vital to a remote user community.

Our main criticism of the SUMEX resource regards the documentation for the EMYCIN software. As mentioned, we have frequently found need to contact the MYCIN staff with questions about the software. While they have always been generous with their time and good humored in answering our questions, most of these contacts would have been unnecessary if adequate documentation existed. Current industry practice is to generate about two pages of documentation per page of code listing. The quantity of documentation for MYCIN and EMYCIN combined falls well below this amount. We note, in this regard, that modern programming practice dictates the creation of most of the documentation before any of the code is written.

III. Research Plans (August 1978 - July 1981)

A. Long Range Project Goals and Plans.

1. Evaluation of the Psychopharmacology Advisor.

When the performance of the Psychopharmacology Advisor approaches an optimal level in the judgment of the Principal Investigators and the Advisory Panels, a formal evaluation will be performed. Elaborate plans have been made for three types of evaluation: as a simulation of the Principal Investigator; as a national expert; and as an actual psychopharmacology advisor. In each evaluation the system will be tested on two sets of cases: one which represents the population of patients likely to be encountered in practice, thereby measuring whether HEADMED can do well what it must do most often; and one which represents unusual or exceedingly complicated cases, thereby measuring whether the program can do well in situations where usual practices may not suffice. Details of the evaluation plans are available upon request.

We will also evaluate the EMYCIN formalism, regarding both its inherent properties as a consulting algorithm and its appropriateness for the domain of clinical psychopharmacology. If the results of the evaluation of either EMYCIN's formal properties or EMYCIN's applicability to the clinical psychopharmacological domain are less than ideal, which is almost certain to be the case, then we plan to construct a new system, HEADMED-II, incorporating what we have learned into hardware and software technologies then available.

B. Justification and Requirements for Continued SUMEX Use.

As mentioned in the preceding paragraph, we consider use of the EMYCIN software as integral to our project, at least, for the next two years. While we previously contemplated adapting the EMYCIN software to our local DECsystem-10 environment, we have since discovered that the conversion process would be tantamount to writing a new system. Therefore, at least until we have learned enough about the domain of clinical psychopharmacology to know how to supersede the EMYCIN formalism, we will have a continuing need for the SUMEX resource.

Our typical work pattern over the past year has been to hold intense working sessions of approximately a month followed by 2 to 3 month periods of only occasional system use. During our working sessions, our share of processor time has been adequate. During the coming year, we expect our working sessions to occur about every other month, resulting in, approximately, a doubling of our need for processor time.

At least part of this increased need for processor time can be reduced by increased disk space. Currently, the 300 page allocation on our main account is not sufficient to hold a SAV file for a loaded system; as a consequence, we use substantial CPU resources in repeatedly reloading our system. An increase in disk allocation to 1000 blocks would reduce this use.

Another need we have for increased resources results from the large number of short-term staff on our project. Since we have residents and medical students for only 2 or 3 month periods, we need to get them up to speed as quickly as possible. The ability to establish our own on-line bulletin board, with sufficient disk space (150 pages) to support it, would be useful in this regard.

C. Our Needs and Plans for Other Computational Resources, beyond SUMEX/AIM.

Our only immediate need for other computational resources beyond SUMEX/AIM is for local, high-speed printing, preferably combined with local file storage. Our current slow-speed printing is unsuitable for listings of large rule sets or of system code. The planned acquisition of a 1200 baud printing terminal may substantially reduce the problem.

Our future plans will depend greatly on the outcome of our current effort. If the EMYCIN formalism proves suitable for our domain, we may find the conversion effort sufficiently worthwhile to transport EMYCIN to our local environment. If we discover that a major redesign is needed, we will make our future computing plans in light of that design.

D. Recommendations for Future Community and Resource Development.

We suspect that the documentation problem exemplified in the EMYCIN software also exists in other SUMEX projects. If software developed for these projects is usable only by the original authors, then much of the impact of the work will be lost. We therefore suggest an increased emphasis on system design and documentation tools. Something along the line of the Programmer's Workbench concept, which links together global design, local design, and code, would probably be useful.

4.3.3 ORGAN CULTURE PROJECT

Application of Computer Science to Organ Culture

Professor Robert K. Lindsay and Dr. Maija Kibens
The University of Michigan, Ann Arbor

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

The goal of this research project is to develop, using artificial intelligence techniques, a programmed model of organ culture that will assist histologists in the design and interpretation of their experiments.

In organ cultivation, one strives to create an environment outside of a living organism that will preserve the physiological and structural relationships among tissues of an organ fragment (explant). Culturing organ explants is an important technique in biomedical research on disease processes, nutrition, physiology, and medicinal effects. Experiments performed with organ explants have the advantages of being safer, quicker, less costly, and more easily interpreted than in vivo experiments. In contrast with cell culture, organ culture provides a more realistic environment for studying processes in healthy and diseased tissues and can provide information on the interactions between tissue types. It has been shown to be particularly helpful in studying diseases that are specific to one organ or that manifest themselves differently in different organs, such as cancer.

Organ cultivation involves techniques for obtaining and preparing explants, creating media, and assessing the condition of the explants at various time intervals. The first of these techniques involves primarily adeptness at physical and manipulative skills and is outside the scope of this project. We would, however, like to develop a computer model of the cultivation and evaluation processes. This model would be able to represent the knowledge which histologists have concerning the biochemical processes of life, the optimal structure of cells within an organ, the components of different culture media, staining and fixation techniques for microscope viewing, and the format of experimental designs. The desired computer system should be able to interpret microscope pictures of organ explants and store a representation of their appearance. It should also be able to interpret natural language descriptions of organ explants and store this information in the same internal format as that for images. The system should be able to access this knowledge base to answer a researcher's inquiries about particular explants and make comparisons between explants. It should also be able to interpret the relative success of different cultivation techniques and guide a researcher in the planning of future experiments.